



**NORTHCENTRAL UNIVERSITY
ASSIGNMENT COVER SHEET**

Student: **Michael Higley-Vance**

THIS FORM MUST BE COMPLETELY FILLED IN

Follow these procedures: If requested by your instructor, please include an assignment cover sheet. This will become the first page of your assignment. In addition, your assignment header should include your last name, first initial, course code, dash, and assignment number. This should be left justified, with the page number right justified. For example:

DoeJXXX0000-1

1

Save a copy of your assignments: You may need to re-submit an assignment at your instructor's request. Make sure you save your files in accessible location.

Academic integrity: All work submitted in each course must be your own original work. This includes all assignments, exams, term papers, and other projects required by your instructor. Knowingly submitting another person's work as your own, without properly citing the source of the work, is considered plagiarism. This will result in an unsatisfactory grade for the work submitted or for the entire course. It may also result in academic dismissal from the University.

EDU7006-8

Dr. Rebecca Watts

Quantitative Research Design

**Activity #2b: Exploring Inferential
Statistics and Their Discontents**

Comments:

Faculty Use Only

Michael, thanks for submitting this part of the week 2 assignment. I see that you answered all of the questions for this assignment. Be sure to answer all parts of the assignment. I also want to encourage you to use the Jackson textbook in defining terms and answering questions. The questions and assignments follow the reading of that text very closely and I encourage you to rely more heavily on that source in answering your questions. I also encourage you to use the Trochim source. It seems that you are trying to use many sources and, in doing so, you may not get a very thorough understanding of the terms and concepts. So, I would read the information in the Jackson textbook for sure and rely more heavily on that

[source and the Trochim source to help you in answering the questions. Please let me know if you have any questions about my comments.](#)

[Score = 100](#)

<Faculty Name>

<Grade Earned>

<Writing Score>

<Date Graded>

Numerical Points	Letter Grade	Descriptor	Explanation
100 - 94	A	Excellent	Completes all required parts of the assignment, demonstrates deep understanding of materials, uses very clear and effective expression appropriate to scholarly writing, and has very few or no errors in grammar, mechanics, and APA formatting.
93-90	A-		
89-87	B+	Good	Completes all or most required parts of the assignment, demonstrates good understanding of readings, uses mostly clear and effective expression appropriate to scholarly writing, and has few errors in grammar, mechanics, and APA formatting.
86-83	B		
82-80	B-	Fair	Completes most required parts of the assignment, demonstrates some understanding of readings, and writing is somewhat clear, effective , and scholarly, and has some errors in grammar, mechanics, and APA formatting.
79-77	C+		
76-73	C	Poor	Completes some required parts of the assignment, demonstrates some understanding of readings, and writing is difficult to understand and unscholarly and has several errors in grammar, mechanics, and APA formatting.
72-0	F	Unacceptable	Completes few required parts of the assignment, demonstrates little understanding of readings, and

			writing is difficult to understand and unscholarly and has many errors in grammar, mechanics, and APA formatting.
--	--	--	--

Answer the following questions:

- What does $p = .05$ mean? What are some misconceptions about the meaning of $p = .05$? Why are they wrong? Should all research adhere to the $p = .05$ standard for significance? Why or why not?
- Compare and contrast the concepts of effect size and statistical significance.
- What is the difference between a statistically significant result and a clinically or “real world” significant result? Give examples of both.
- What is NHST? Describe the assumptions of the model.
- Describe and explain three criticisms of NHST.
- Describe and explain two alternatives to NHST. What do their proponents consider to be their advantages?
- Which type of analysis would best answer the research question you stated in Activity 1? Justify your answer.

Length: 5-7 pages not including title and reference pages. References: Minimum of 3-5 scholarly resources. Your paper should demonstrate thoughtful consideration of the ideas and concepts that are presented in the course and provide new thoughts and insights relating directly to this topic. Your paper should reflect scholarly writing and current APA standards. Review [APA Form and Style](#). Be sure to adhere to Northcentral University's Academic Integrity Policy. View the [Northcentral Academic Integrity Tutorial](#) to refresh your knowledge of how to achieve academic integrity.

Exploring Inferential Statistics and Their Discontents

1. What does $p = .05$ mean? The P value is the estimated probability of rejecting the null hypothesis (H_0) of a study when that hypothesis is true. According to Jackson (2012) “probability refers to the number of ways a particular outcome can occur divided by the total number of outcomes” (p. 173). The estimated probability is compared to the significance level. If $P = .05$, the null hypothesis only has a 5% chance of being true. If the p-value exceeds the established level, or alpha, then the alternative hypothesis (H_a) is accepted (Trochim & Donnelly, 2008). If it does not the null hypothesis is accepted instead. What are some misconceptions about the meaning of $p = .05$? Why are they wrong? If the probability value is less than the indicated significance level researchers must conclude that the alternative hypothesis is true, however if the p-value exceeds the significance level then the alternative hypothesis is rejected leaving only the null hypothesis (Trochim & Donnelly, 2008). This very idea seems easy to follow however, 48% of researchers fail to interpret the p-value correctly (Goodman, 2008). Goodman (2012) suggested that the statistician who introduced the use of probability into the research field, Sir Ronald Fisher, could barely explain the inferential standpoint himself. Several misconceptions about the significance of considering the p-value have been

plaguing the scientific world since the 1940's (Goodman, 2008). The most common misconception is that the null hypothesis only has a 5% chance of being true. This misconception perpetuates the idea that only data can develop conclusions to research studies (Goodman, 2008). Some authors have argued against this point stating that in fact researchers are more likely to conclude a false find than to correctly conclude that it does not (Simmons, Nelson, & Simonsohn, 2011). Another misconception is that $p = .05$ is the same as $p \leq .05$. There is a huge difference between each expression but because both use the same number (5%) it is almost impossible to explain without the use of an alternative mathematical equation (Goodman, 2008). Should all research adhere to the $p = .05$ standard for significance? Why or why not? Not all research should adhere to the p-value standard of .05 (5%) because it is not appropriate for all research to be measured using this standard. It is appropriate for some research to use a different significance level depending on the research design, question, statistical practice, and argument being made (Simmons, Nelson, & Simonsohn, 2011).

2. Compare and contrast the concepts of effect size and statistical significance. Population effect sizes are typically estimated by examining representative samples of a larger population or group (Jackson, 2012). The goal of minimizing the measurement and sampling error produces more accurate effect size estimates, which provides a better scientific understanding of real world effects (Trochim & Donnelly, 2008). The significance or alpha level of any test result is determined by the probability of getting a result, and the outcome is a conditional probability or p-value. If the significance level falls below the accepted standard, the result is concluded to be statistically significant (Trochim & Donnelly, 2008).

3. What is the difference between a statistically significant result and a clinically or “real world” significant result? Give examples of both. The standard for statistical significance is used in hypothesis testing (Trochim & Donnelly, 2008). When research is found to be statistically significant or non-significant, the null hypothesis is rejected, however, researchers are unable to prove the hypothesis is true or false (Monterde-i-Bort, Frias-Navarro, & Pascual-Llobell, 2010; Simmons, Nelson, & Simonsohn, 2011). Statistical significance implies only to the probability that the research results obtained were not due to chance (Trochim & Donnelly, 2008). Clinical research is the practical importance of a genuine or noticeable treatment effect (Trochim & Donnelly, 2008). Clinical research provides evidence for most therapeutic advances, and that evidence, is expressed in a statistical format (Monterde-i-Bort, Frias-Navarro, & Pascual-Llobell, 2010). Clinicians struggle with the interpretation of statistical results to make inferences about a population making clinical judgments about individual patient health and care (Mariani & Pêgo-Fernandes, 2014). For example, if a researcher wanted to know the average number of cigarette packs Tennesseans smoke in a month compared to the average Floridian. The researcher could sample a small number of people who smoke from each state providing results that could be skewed or interpreted as chance. However, if the researcher took a clinical sampling from a larger control group or population the results are likely to be more statistically significant, plausible, and accurate (Man-Son-Hing et al., 2002).
4. What is NHST? Null hypothesis significance testing (NHST) is the most accepted approach to a statistical inference in quantitative research (Gliner, Leech, & Morgan, 2002). Decisions made using NHST are not formed from the probability of a specific

outcome but instead from all possible outcomes and are highly controversial (Levine, Weber, Hullett, Hee Sun, & Lindsey, 2008). Describe the assumptions of the model. Researchers use different hypothesis tests to make different assumptions about the variable being sampled. It is important to consider these assumptions when choosing a test and interpreting the results. For example, the z-test and the t-test both assume that the population is randomly selected, and the data are normally distributed (Gliner, Leech, & Morgan, 2002).

5. Describe and explain three criticisms of NHST. A common criticism of NHST is the implication that statistical significance also implies practical significance. Researchers often interpret a statistically significant difference as a difference that has clinical implications. The implication of such a finding may be sufficient enough to be determined to have practical significance when indeed it should not (Gliner, Leech, & Morgan, 2002). Another criticism states that statistical significance testing is based upon a central misunderstanding of rational inference and is seldom appropriate to the aims of scientific research (Levine et al., 2008). Finally, some researchers suggest that testing for statistical significance has become a tradition rather than having a necessary purpose to the study and therefore is prone to error (Levine et al., 2008).
6. Describe and explain two alternatives to NHST. What do their proponents consider to be their advantages? The first alternative to NHST is the use of a statistic prep, which estimates the probability of replicating an effect. It captures traditional publication criteria, provides all of the information now used in evaluating research, and avoids many criticisms of traditional statistical inference (Killeen, 2005). Another alternative to NHST is the use of confidence intervals, which are another means of making sense of

data. A confidence interval provides a standard for assessing the probability that an effect lies within a particular range of values. Furthermore, a confidence interval tells researchers more than a null hypothesis significance test (Denis, 2003).

An alternative to NHST is to use confidence intervals (Carver, 1978; Gelo et al., 2008; Schmidt, 2010). Confidence intervals provide additional information than significance tests and can be just as objective (Carver, 1978; Gelo et al., 2008; Schmidt, 2010). Confidence intervals provide a range of values within the population parameters is likely to lie and are especially useful in determining real world significance of the research results (Carver, 1978; Gelo et al., 2008; Schmidt, 2010). A second alternative to NHST is to report the effect size in research results (Carver, 1978; Jackson, 2012; Schmidt, 2010). Effect size summarizes data by identifying the significance of an association. Researcher that reports effect size in the study can make claims of the strength estimation of an association regarding the practical utility of the results (Carver, 1978; Jackson, 2012; Schmidt, 2010).

7. Which type of analysis would best answer the research question you stated in Activity 1?

Justify your answer. A confidence interval because such a measure would inform the researcher as to the probability that a given range includes the given effect size. For instance, a 92% confidence interval for the effect size of 0.6 would imply that we are 92% sure that the true effect size lies between 0.3 and 0.8. Confidence intervals represent a way of quantifying the degree of certainty, or uncertainty, with regards to the true population parameter (Denis, 2003). Confidence intervals are also a good way to report the results of a study because the provision of intervals is a useful and necessary means of accounting for error, and provides the researcher with a more meaningful answer to the hypothesis (Denis, 2003; (Trochim & Donnelly, 2008).

References

- Denis, J. (2003). Alternative to null hypothesis significance testing. *Theory & Science*. Retrieved from http://theoryandscience.icaap.org/content/vol4.1/02_denis.html
- Gliner, J. A., Leech, N. L., & Morgan, G. A. (2002). Problems with null hypothesis significance testing (NHST): What do the textbooks say?. *Journal Of Experimental Education*, 71(1), p. 83. Retrieved from <http://proxy1.ncu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=7217331&site=eds-live>
- Goodman, S. (2008). A dirty dozen: twelve p-value misconceptions. *Seminars In Hematology*, 45(3), 135-140. Retrieved from <http://www.perfendo.org/docs/BayesProbability/twelvePvaluemisconceptions.pdf>
- Jackson, S. L. (2012). *Research methods and statistics: A critical thinking approach*. Belmont, CA: Wadsworth Cengage Learning.
- Killeen, P. R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science (Wiley-Blackwell)*, 16(5), 345-353. doi:10.1111/j.0956-7976.2005.01538.x
- Laerd Statistics. (2013). Independent T-Test - An introduction to when to use this test. Retrieved from <https://statistics.laerd.com/statistical-guides/independent-t-test-statistical-guide.php>
- Levine, T. R., Weber, R., Hullett, C., Hee Sun, P., & Lindsey, L. (2008). A critical assessment of null hypothesis significance testing in quantitative communication research. *Human Communication Research*, 34(2), pp. 171-187. doi:10.1111/j.1468-2958.2008.00317.x
- Man-Son-Hing, M., Laupacis, A., O'Rourke, K., Molnar, F., Mahon, J., Chan, K., & Wells, G. (2002). Determination of the clinical importance of study results. *Journal Of General Internal Medicine*, 17(6), pp. 469-476.
- Mariani, A., & Pêgo-Fernandes, P. (2014). Statistical significance and clinical significance. *São Paulo Medical Journal = Revista Paulista De Medicina*, 132(2), pp. 71-72. Retrieved from <http://proxy1.ncu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=mnh&AN=24714985&site=eds-live>
- Monterde-i-Bort, H., Frias-Navarro, D., & Pascual-Llobell, J. (2010). Uses and abuses of statistical significance tests and other statistical resources: A comparative study. *European Journal Of Psychology Of Education*, 25(4), pp. 429-447. Retrieved from <http://proxy1.ncu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ902644&site=eds-live>
- Schmidt, F. (2010). Detecting and correcting the lies that data tell. Retrieved from <http://pps.sagepub.com.proxy1.ncu.edu/content/5/3/233.full.pdf%20html>

- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science (Sage Publications Inc.)*, 22(11), pp. 1359-1366.
doi:10.1177/0956797611417632
- Trochim, William M. (October 20, 2006). The Research Methods Knowledge Base, 2nd Edition.
Retrieved from <http://www.socialresearchmethods.net/kb>
- Trochim, W. M., & Donnelly, J. P. (2008). *Research methods knowledge base*. Mason, OH: Atomic Dog/Cengage Learning.